

RUPAGG:
AGGREGATION OF RUP PROJECTIONS

Peter D. Johnson
International Programs Center
Population Division
U.S. Bureau of the Census
December 1994 (revised Dec 1999)

1. Introduction

The RUPAGG program aggregates population projections produced by the RUP program. The program works with integral numbers of population, births, deaths and net numbers of migrants from different areas to arrive at a total. These population values must be stored in intermediate files by the RUP runs. This is accomplished by including CODE records in the RUP runs (see Eduardo Arriaga and Associates, 1994, Population Analysis with Microcomputers, vol. II, p. 332).

2. General Structure of Input to RUPAGG

The input to RUPAGG consists of two types of records (just like RUP): parameter records and data records. The parameter records can serve three functions:

- (1) describe certain aspects of the projection (e.g. the TOT record indicates that the records that follow apply to the total area projection),
- (2) define parameters of the projection (e.g. the PROJ record indicates the final year of the projection), or
- (3) introduce certain data records.

The parameter records all follow the same fixed format as RUP. The parameter records that are recognized by the program are shown in table 1.

The data records contain the data of a repetitive nature, and allow for formats specific to the information they contain (e.g. character data for the title records and 5-column fields of numeric data for the special age groups). The formats of the data records are identical to the same records in RUP. The format for the records in RUPAGG that are not part of RUP are shown in section 2.1.2 to 2.1.4.

Table 1. RUPAGG Parameter Records and Their Functions

Type	Description
TITL	Precedes records with descriptive information to be printed on each page of output.
N5	Specifies the number of 5-year age groups to be used in the aggregation.
N	Specifies the number of single years of age to be used in the aggregation.
SPAG	Specifies the special age groups for which population data is to be printed.
PROJ	Specifies the final year of the projection.
AREA	Names the aggregated area.
CODE	Specifies a code number to be associated with the aggregated area and creates an intermediate file with the aggregated data.
TOT	Indicates the area name is "T O T A L".
BASE	Indicates desired base year of aggregation.
INPF	Indicates the code number of each area and the name of the intermediate file where the data is stored.
RUM	Indicates internal migration data is present.
OUTP	Controls full-page output.
OPOP	Controls special population output.
OMX	Controls age-sex-specific central death rate output.
ODTH	Controls output of deaths by age and sex.
OBTH	Controls output of births by age of mother.
END	Indicates the end of the projection inputs.
NOTE	Allows the inclusion of descriptive notes which are printed only as encountered during input.
*	Alternate note format.
EDIT	Allows scanning of parameter and data records without projecting.

2.1. Parameter Records

2.1.1. Parameter Record Rules

The parameter records and any associated data records DO NOT need to follow certain rules regarding where they are located in the input to RUPAGG.

Only One Required: TITL, [N5, N], PROJ, [AREA, TOT], BASE

One or more required: INPF

```
Optional:      EDIT, SPAG, CODE, OUTP, OMX, OPOP, ODTN, OBTH,
                END, RUM
```

Notes:

[A, B] Choose only one parameter record, A or B.

The NOTE parameter record and associated notes or the * format of the notes can be placed at any location in the run where a parameter record is expected.

2.1.2. The BASE Record: Initial Projection Year

The BASE record is used to indicate the beginning year of the projection. One BASE record with a *ybeg* value greater than or equal to 0 is required for each run. The *ybeg* value must be greater than or equal to the beginning year of each of the projections being aggregated (but the individual projections do not need to begin with the same year).

```
*-----*
*          10           20           30           40           50           60
*-----|-----|-----|-----|-----|-----|
BASE   ybeg
*-----*
S))))))0)))))))))0)))))))))0)))))))))
Record * Columns * Field * Definition
S))))))2)))))))))2)))))))))2)))))))))
    1         1-4      BASE     Indicates this is a BASE record.
                                7-10       ybeg      Beginning year of the projection.
S)))))))))
Example: Specify the beginning year of the aggregation.
*-----*
*          10           20           30           40           50           60
*-----|-----|-----|-----|-----|
BASE   1995
*-----*
The populations should be aggregated starting in 1995.
```

2.1.3. The INPF Record: Area Reference Number for Input Intermediate Files

The INPF record specifies the area reference number to help identify the data and the data record that follows contains the name of intermediate file.

	10	20	30	40	50	60
INPF	arnum					
filename						
Record	* Columns	* Field	* Definition			
1	1-4	INPF	Indicates this is a <u>INPF</u> record.			
	5-14	blank				
	15-20	arnum	Area reference number. This number must be right-justified. A negative number indicates that this area should be subtracted from the aggregation.			
2	1-80	filename	Name of the intermediate file with the data corresponding to area arnum. If no path or drive is specified, the default drive is assumed.			

S)))))))))))))
Example 1: Aggregate urban and rural areas

```

*          10          20          30          40          50          60
*          |          |          |          |          |          |
...
TOT
INPF          1
URBAN.IO1
INPF          2
RURAL.IO1
*-----
The first area, urban, denoted area 1, and the second area, rural,
denoted area 2, are to be aggregated. Based on the intermediate
file names, assuming the default naming conventions, we can conclude
that the urban data was generated by an input file called URBAN.IN
and the rural input file was RURAL.IN. If both intermediate files
were created in one RUP run, called for example RUP.IN, then the
first intermediate file created would be named RUP.IO1 and the second
would be called RUP.IO2.

```

Example 2: Compute rural as a residual

* ..	
TOT	
INPF	1
TOTAL.IO1	
INPF	-2
URBAN.IO1	

The first area is the total, denoted area 1, while the second area, urban, is denoted area 2, and will be subtracted from the total. Based on the intermediate file names, assuming the default naming conventions, we can conclude that the urban data was generated by an input file called URBAN.IN and the total input file was TOTAL.IN.

2.1.4. The RUM Record: Internal Migration is Present

The RUM record is used to indicate that one or more of the intermediate files contain internal migration data. If this record is not present, the program will ignore internal migration data stored in the intermediate files.

```
*-----*
*          10          20          30          40          50          60
*-----|-----|-----|-----|-----|-----|
RUM
*-----*
S))))))0)))))))))0))))))0)))))))))
Record * Columns * Field * Definition
S))))))2)))))))))2))))))2)))))))))
   1         1-4        RUM      Indicates this is a RUM record.
S)))))))))
Example:  Indicate presence of internal migration data
*-----*
*          10          20          30          40          50          60
*-----|-----|-----|-----|-----|
RUM
*-----*
The populations being aggregated contain internal migration.
```

3. Sample input file

```
TITL                1
TOTAL OF THE TWO RUPTSTC OUTPUTS (IDENTICAL AREAS)
BASE  1985
PROJ  1990
TOT
INPF                1
RUPTSTC.IO1
INPF                2
RUPTSTC.IO2
RUM
OUTP  1985
OMX   1990
END
```

Notes:

1. If no drive path is entered in the file name record (following the INPF record), the program will look for the file on the default drive/path.
2. The number of areas that can be aggregated is limited to 300. To aggregate more areas, include a CODE record in the RUPAGG input file to indicate you want the program to create an intermediate file of the aggregation. This intermediate file can then be used in another aggregation run.